

Exploring the Implementation of Automatic Text Summarization in Online Learning Setting

Agus Wedi¹, Saida Ulfa^{1,*}, Ashar Pakkawaru², Rex Bringula³

¹ Department of Educational Technology, Faculty of Education, Universitas Negeri Malang, Malang, Indonesia

² Department of Electrical Engineering, Faculty of Engineering, Universitas Negeri Malang, Malang, Indonesia

³ College of Computer Studies and Systems, University of The East, Manila, Philippine

* Corresponding author. Email: saida.ulfa.ftp@um.ac.id

ABSTRACT

Automatic text summarization is a technique of text mining which is a field of natural language processing. It summarizes a long text file or document automatically into several sentences without losing its essence. The use of automatic text summarization in the online learning context has not been widely studied. The purpose of this study is to explore the implementation of automatic text summarization in the online learning context from the year 2010 to 2020. This study uses a literature review process and the results of this study provide an overview of the use of text summarization in online learning context.

Keywords: automatic text summarization, online learning, text mining

1. INTRODUCTION

Text mining is part of an artificial intelligence (AI) which can be used to explore data in the form of text. Text mining is in the field of computational linguistics and is content mining where the process of extracting or mining information is in the form of text that contains unstructured information (Feldman & Sanger, 2007). Text mining is a system that analyzes large amounts of natural language text and detects lexical or linguistic patterns in an effort to extract them to produce useful information (Sebastiani, 2002).

It condenses the text while maintaining the essence of the original text. Text summarization can be done manually, but it is very ineffective and inefficient, especially when regarding a lot of data and sources where it is not possible to perform the task manually. Therefore, automatic text summarization is needed to in order to process larger amounts of text data.

In fact, text summarization research is not new and was started in the 1950s (Luhn, 1958), and the techniques are still being continuously developed. The research related to text mining has increased in recent years (Anand & Wagh, 2019), where various types of algorithms are applied to produce an optimal text analysis.

However, it is not yet widely used in the world of education. This is in spite of the fact that using text mining can increase the effectiveness and efficiency in the teaching and learning process. The purpose of this study was to explore the use of text summarization in an online learning context using literature review research.

2. LITERATURE REVIEW

a. Text Mining in Online Learning

Text mining or knowledge discovery from text (KDT) was first introduced by Feldman and Dagan (1995). Text Mining is the process of searching for information from a set of documents using a set of tools for analysis (Feldman & Sanger, 2007). Text Mining combines data mining techniques, machine learning, natural language processing, information retrieval, and knowledge management. The useful information is obtained by identifying and exploring the text. The analysis process in Text Mining includes Information Retrieval, Text Analysis, Information Extraction, Text Summarization, Clustering, Categorization, Visualization, Machine Learning, and Data Mining (Tan, 1999).

b. Automatic Text Summarization

Automatic text summarization is part of Natural Language Processing (NLP) and is the process of shortening or summarizing a text source or a collection of text documents or paragraphs while maintaining the main information content. The main objective of Text Summarization is to create a reduced version of the text while preserving the essential information (Rajasekaran & Varalakshmi, 2018). In general, automatic text summarization methods are classified into two main categories: (1) extractive methods namely selecting the main important sentiments from the document using different statistical methods; and (2) the abstractive method

creates a semantic representation of the input text to generate a summary.

After the text has been processed, important sentences are identified in the document which can be considered further for the summarization process. The text processing is conducted through summation using several different features to determine the weight of each sentence in the document. The score for the next sentence is calculated based on the linear combination of these derived weights. Once sentence scores are calculated, sentences are ranked in descending order based on their score - from highest score on top to lowest at bottom. The sentence with the highest score is taken for summary production.

3. METHOD

In this study a systematic literature review was used aimed to identify, evaluate and analyze the results of the review based on research objectives and a certain question. In this study, the researchers proposed a research question “how is the use of text summarization in online learning? To answer the research question, a literature review process was carried out using three stages as shown in Figure 1.

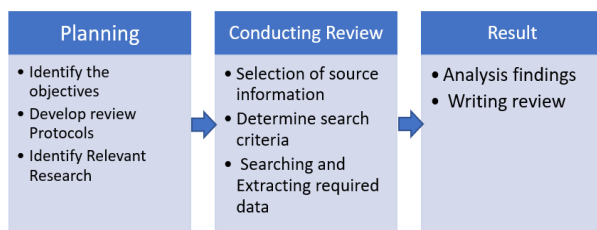


Figure 1 Literature Review Process

The stages are: (1) planning stage consisting of identifying objectives, developing review protocols and identifying relevant research; (2) conducting review stage consisting of a selection of information resources, determining search criteria (inclusion and exclusion criteria), searching and extracting required data; and (3) result stage which is the stage of analyzing findings and writing reviews. Table 1 shows the criteria used to filter the papers.

The papers that fall into the criteria are articles published through a review process so that articles published in conference proceedings and journals are selected. In addition, the article must be published between year 2010 and 2020. To facilitate analysis of papers reviews, the researchers required the papers to be written in English. There are two exclusion points for the criteria in this paper, which are that the studies do not have experimental results and studies do not have unclear summarization text strategy and data sets.

4. RESULTS AND DISCUSSION

Many studies on automatic text summarization have been carried out. However, there are only few studies which relate to the implementation of text summarization

implemented in education. The search strategy used the procedure, with the first stage being “select digital libraries”. Researchers in this study used the Google Scholar database as consideration that many reputable international journals use the Google Scholar index as well.

Table 1 Inclusion and Exclusion Criteria

Criteria	Explanation
Inclusion	<ol style="list-style-type: none"> 1. Studies that published in conference proceeding and journal 2. Studies which published in 2010 – 2020 3. Studies which are written in English
Exclusion	<ol style="list-style-type: none"> 1. Studies do not have experimental result 2. Studies do not have unclear summarization text strategy and data set 3. Studies do not discuss about implementation of summarization text in online learning environment

The second stage is “define keywords” using the keyword “implementation of text summarization in online learning”. The next stage is the search process by following the predetermined requirement being that the papers should be published between years 2010 and 2020. This search process yielded 648 papers. The following stages are “Filtering titles and abstract” of the 648 obtained papers, where filtering was carried out using several inclusion and exclusion criteria listed in Table 1. The following is a list of papers found based on the search results.

This research found 10 papers focusing on implementation of text summarization in online learning. In order to answer the research question posed in this paper, the obtained papers listed in Table 2 were explored. The following is the explanation of the research findings. One paper discusses implementation of automatic text summarization in mobile learning context (Yang, Chen, Kinshuk, & Sutinen, 2013) and the remaining papers discuss implementation of text summarization in online learning context.

The implementation of text summarization in online learning can be seen in its use in online forum discussions. There are four papers which discuss this. Hsio and Awasthi (2015) made a prototype of a visual analytics interface to present the semantic of online discussion forum for helping online learners to examine the post content by viewing text summarizations. Tarnpradab, Liu, and Hua (2018) made thread summarizations in an online discussion forum to help online learners find information more easily. In addition, the study of Gottipati, Shankararaman, and Ramesh (2019) extracted insights from online discussion forums through generation of topic based summaries.

The results of this text summarization are used by students and teachers to focus more on the topics discussed in a forum. In online discussion forums, there are lots of posts and threads. These three studies are aimed at exploring ideas, insights, and hidden knowledge in a discussion forum using text summarization techniques. Another use of text summarization in online forums can be seen from the study of Siu (2016), where he developed an approach to assess the quality of posts in online discussion forums to avoid clutter within online learners.

Table 2 List of Papers

Publications	Titles
Guangbing Yang; Nian-Shing Chen; Kinshuk; Erkki Sutinen Computers & Education, Volume 68, October 2013, Pages 233-243	The effectiveness of automatic text summarization in mobile learning contexts
I-Han Hsiao; Piyush Awasthi LAK '15: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge. 2015	Topic facet modeling: semantic visual analytics for online discussion forums
Chapman Siu. 2016. School of Computer Science Graduate Student Publications	Duplicate Question Detection Using Online Learning
Ferry Pramudianto; Tarun Chhabra; Edward F. Gehringer; Christopher Maynard. http://ceur-ws.org/Vol-1633/ws1-paper12.pdf . 2016	Assessing the Quality of Automatic Summarization for Peer Review in Education
Sansiri Tarnpradab, Fei Liu, Kien A. Hua. Computation and Language. https://arxiv.org/abs/1805.10390	Toward Extractive Summarization of Online Forum Discussions via Hierarchical Attention Networks
Derek Miller. Computer Science > Computation and Language .Cornell University. https://arxiv.org/abs/1906.04165	Leveraging BERT for Extractive Text Summarization on Lectures
Fábio Bif Goulartea; Silvia Modesto Nassar; Renato Fileto; Horacio Saggion Expert Systems with Applications Volume 115, January 2019, Pages 264-275	A text summarization method based on fuzzy rules and applicable to automated assessment
Swapna Gottipati ; Venky Shankararaman ; Renjini Ramesh 2019 IEEE Frontiers in Education Conference (FIE). 2019	Topic Summary: A Tool for Analyzing Class Discussion Forums using Topic Based Summarizations
Luca Cagliero; Laura Farinetti; Elena Baralis. IEEE Access, vol. 7, pp. 22729-22739	Recommending Personalized Summaries of Teaching Materials
Yusra AlRoshdi ; Mohammed AlBadawi; Mohammed Sarrab 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS)	Framework for Socializing Learning Content

According to Table 2, there are three papers that discuss the use of text summarization on learning content. As we know, learning content is one of the essential components in online learning. Learning content can be presented in various formats such as text, pdf, images, videos, links to learning resources and audio. The enormous amount of learning content in a learning management system (LMS) can be overwhelming and difficult to find manually.

Therefore, research related to text mining, in this case text summarization is needed, in order to increase the effectiveness and efficiency in utilizing learning content. The research conducted by Miller (2019) focuses on developing tools that can summarize lecture content automatically, making it easier for students to read summaries of lectures that had been presented by the instructor, presented in shorter text or documents. Cagliero, Farinetti, and Baralis (2019) conducted a study on recommendations of short documents or study content that suit the needs of learning participants.

A research conducted by AlRoshdi, AlBadawi, and Sarrab (2020) was to summarize learning content on the e-learning platform (learning management system) and then disseminate it through a social network platform. They chose social media with the consideration that online learners currently spend more time using social network platforms than using e-learning platforms. These three studies were all aimed at improving learning experience and increasing learning outcomes. There are two studies that discuss assessment using the automatic text summarization technique. Firstly, in research conducted by Pramudianto, Chhabra, and Gehring (2016) on assessment through peer review apply the text summarization technique on the results of feedback from peer reviews conducted by students.

The purpose of this study was to make it easier for students to read a summary of their friends' reviews, and by reading a shorter text, it motivated the learners to read the reviews. On the other hand, often what occurs is when someone finds feedback in the form of a long text, they overlook or ignore the feedback. The research which related to assessment was also conducted by Goulartea, et al. (2019). In their research fuzzy logic was used associated with text summarization on a computer-assisted assessment.

5. CONCLUSIONS

There are only few studies that applied automatic text summarization in the world of education, especially in the online learning context. Based on the literature review findings based on the publication within the years 2010 and 2020; ten papers were found that discuss the implementation of text summarization in online learning context. Based on the specific discussed topics of the 10 papers, it can be categorized into several categories; the use of text summarization in: (1) mobile learning; (2) online discussion forums; (3) learning content; and (4) assessments.

REFERENCES

- [1] AlRoshdi, Y., AlBadawi, M., & Sarrab, M. (2020). Framework for socializing learning content. 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS), (pp. 1-5). Riyadh.
- [2] Anand, D., & Wagh, R. (2019). Effective deep learning approaches for summarization of legal texts. Journal of King Saud University - Computer and Information Science.

- [3] Baturay, M. H. (2015). An overview of the world of MOOCs. *Procedia - Social and Behavioral Sciences* 174 (2015), 427 – 433
- [4] Cagliero , L., Farinetti, L., & Baralis, E. (2019). Recommending Personalized Summaries of Teaching Content. *IEEE Access*, vol. 7, 22729-22739.
- [5] Feldman, R., & Sangar, J. (2007). *The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data*, Newyork
- [6] Feldman, R. and Dagan, I. (1995). KDT-Knowledge discovery in texts. In *Proceedings of the First International Conference on Knowledge Discovery (KDD-95)*.
- [7] Gottipati, S., Shankaraman, V., & Ramesh, R. (2019). Topic Summary: A tool for analyzing class discussion forums using topic based summarization. 2019 IEEE Frontiers in Education Conference 49th FIE, (pp. 1-10).
- [8] Goulartea, F. B., Nassar, S. M., Fileto, R., & Saggion, H. (2019). A text summarization method based on fuzzy rules and applicable to automated assessment. *Expert Systems with Applications*. Vol. 115, 264-275.
- [9] Hsio, I.-H., & Awasthi, P. (2015.). Topic facet modeling: semantic visual analytics for online discussion. *LAK '15: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, (pp. 231–235). <https://doi.org/10.1145/2723576.2723613>.
- [10] Luhn, H.P.(1958). The automatic creation of literature abstracts', *IBM Journal of research and development*, 1958, 2, (2), pp. 159-165
- [11] Luo, W., Liu, F., Liu, Z., & Litman, D. (2016). Automatic Summarization of Student Course Feedback. *Proceedings of NAACL-HLT 2016*, (pp. s 80–85). San Diego, California,.
- [12] Miller, D. (2019). Leveraging BERT for Extractive Text Summarization on Lectures. Cornell University.
- [13] Pramudianto, F., Chhabra, T., & Gehring, E. F. (2016). Assessing the Quality of Automatic Summarization for Peer Review in Education. *WT-EDM 2016 :Workshop and Tutorial Proceedings of EDM 2016* (pp. 1-5). <http://ceur-ws.org/Vol-1633/ws1-paper12.pdf>.
- [14] Rajasekaran, A., & Varalakshmi, R. (2018). Review on automatic text summarization. *International Journal of Engineering & Technology* (7), 456-460.
- [15] Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Comput. Surv.*, 34(1):1–47.
- [16] Siu, C. (2016). Duplicate Question Detection Using Online Learning. Retrieved from https://smartech.gatech.edu/bitstream/handle/1853/56078/cs6460_finalproject.pdf
- [17] Tan, A-W. (1999), *Text Mining: The state of art and the challenges*.
- [18] Tampradab, S., Liu , F., & Hua , K. A. (2018). Toward extractive summarization of online forum discussion via hierarchial attention networks. Cornell University.
- [19] Widyassari, A. P., Rustad, S., Shidik, G. F., Noersasongko, E., Syukur, A., Affandy, A., & Setiadi, D.-R. I. (2020). Review of automatic text summarization techniques & methods. *Journal of King Saud University - Computer and Information Sciences*.
- [20] Yang, G. B., Chen, N. -S., Kinshuk, & Sutinen, E. (2013). The effectiveness of automatic text summarization in mobile learning context. *Computers & Education*. Vol 68., 233-243.
- [21] Zheng, L., & et.al. (2014). Emerging Approaches for Supporting Easy, Engage and effective Collaborative Learning. *Journal of King Saud University, Computer and Information Science*.